June 26, 2024

The Honorable Rohit Chopra
Director, Consumer Financial Protection Bureau
1700 G Street NW
Washington, DC 20552

Subject: Urgent Call for Regulatory Clarity on the Need to Search for and Implement Less Discriminatory Algorithms

Dear Honorable Director Chopra,

On behalf of the Consumer Federation of America (CFA)[1] and Consumer Reports (CR),[2] we are writing to address a critical issue that significantly impacts consumers: the urgent need for regulatory clarity and certainty regarding the expectation that financial institutions search for and implement less discriminatory algorithms (LDAs) in credit underwriting and pricing.

In recent years, the adoption of algorithmic decision-making tools by financial institutions, particularly more complex machine learning (ML) models, has surged. While these advancements have the potential to enhance efficiency and advance financial inclusion, there is growing evidence that they can also perpetuate and exacerbate existing and historical biases, leading to discriminatory outcomes that adversely affect marginalized and underserved communities.

When financial institutions fail to search for and implement LDAs, they undermine consumer trust and contravene principles of fairness and equity that are foundational to our financial system and broader society. In some instances, these failings may rise to the level of violating the law.

When considering how to respond, we believe the appropriate course of action for the CFPB is to provide clear guidance on how lenders should search for and implement less discriminatory alternatives when

[1] The Consumer Federation of America (CFA) is an association of non-profit consumer organizations established in 1968 to advance the consumer interest through research, advocacy, and education. Today, more than 250 of these groups participate in the federation and govern it through their representatives on the organization's Board of Directors. As an advocacy organization, CFA seeks pro-consumer policies on a variety of issues before Congress, the White House, federal and state regulatory agencies, state legislatures, and the courts. We communicate and work with public officials to promote beneficial policies, oppose harmful ones, and ensure a balanced debate on issues important to consumers.

[2] Founded in 1936, Consumer Reports (CR) is an independent, non-profit and nonpartisan organization that works with consumers to create a fair and just marketplace. Known for its rigorous testing and ratings of products, CR advocates for laws and company practices that put consumers first. CR is dedicated to amplifying the voices of consumers to promote safety, digital rights, financial fairness, and sustainability. The organization surveys millions of Americans every year, reports extensively on the challenges and opportunities for today's consumers, and provides ad-free content and tools to 6 million members across the U.S.

using algorithms for credit underwriting and pricing, which is currently lacking. Clear guidance is needed to complement supervisory and enforcement action, which by themselves are insufficient to provide necessary clarity for the market. While enforcement has a positive effect on safeguarding the market, it will necessarily permit some harm to occur as a pre-condition of action. <u>Regulatory clarity, combined with guidance on supervisory expectations, will make both supervision and enforcement more effective and can prevent future harm.</u>

Providing carefully-crafted guidance will also counter claims of innocence from financial institutions that do not conduct robust searches for LDAs, while also aiding those who have good intentions but are unaware of proper techniques for testing. Guidance and examples of how to effectively search for and develop LDAs and demonstrating where institutions have already fulfilled the expectations of the CFPB will address both scenarios.

Clear, explicit guidance will also encourage more financial institutions to use these technologies in ways that ultimately lead to more competition in the market. Some lenders are introducing ML models, either of their own making or through a relationship with a vendor, but many remain on the sidelines due to a lack of regulatory clarity. We are concerned that too many "disruptors" will forge ahead with dangerous models, while conversely, many small and medium-sized financial institutions will fall behind, missing the opportunity to leverage ML models to reach underserved consumers.

The risk of harm to consumers due to prolonged resistance to LDA searches calls for moving beyond the status quo. As long as financial institutions do not see irrefutable evidence of expectations to perform statistically effective reviews, many will fail to do so. Nonetheless, we understand the need to proceed with prudence, particularly with a rapidly evolving technology and given the current political and legislative environment. We suggest a set of practical options that can be implemented within the current regulatory framework that can help to make significant progress in addressing these issues.

In this letter, we discuss the following key issues:

- The CFPB has begun issuing guidance on AI-related issues but has yet to provide formal written guidance on algorithmic discrimination, one of the core risks to consumers.
- Existing disparate impact doctrine, specifically the concept of less discriminatory alternatives, can be a powerful tool to address algorithmic discrimination.
- Model multiplicity, a unique aspect of machine learning, greatly expands the feasibility and ability of identifying an LDA.
- The CFPB should utilize strong but expedient instruments to clarify that there is an obligation in the Equal Credit Opportunity Act (ECOA) to mitigate discrimination and to conduct robust searches for LDAs.
- Ongoing guidance should be provided on key operational aspects of searching for LDAs, including the appropriate techniques for conducting a robust LDA search and appropriate fairness metrics.

These steps can be taken without requiring the CFPB to embark on a new rulemaking for Regulation B. The CFPB should clarify that a robust search for LDAs is a regulatory expectation under ECOA, putting in writing in a formal instrument what CFPB staff have already stated orally on multiple occasions.[3] In addition, a variety of other publications should be utilized to elaborate on supervisory expectations

[3] Blower, Brad. "CFPB Puts Lenders & Fintechs on Notice: Their Models Must Search for Less Discriminatory Alternatives or Face Fair Lending Non-Compliance Risk," April 5, 2023. https://ncrc.org/cfpb-puts-lenders-fintechs-on-notice-their-models-must-search-for-less-discriminatory-alternatives-or-face-fair-lending-non-compliance-risk/.

regarding how to search for and implement an LDA as part of compliance management, including supervisory highlights, circulars, and guidance documents.

Guidance on testing for and mitigating discrimination in ML models for lending has been lacking among recent CFPB AI-related policy measures

i. Recent CFPB policy measures on AI-related issues

The CFPB has begun issuing important and concrete guidance on a range of AI-related topics. The CFPB has issued recent statements regarding the use of AI in marketing and requiring that consumers receive explainable answers for adverse credit decisions. It has also provided insights on cutting-edge issues, such as a recent spotlight on risks to consumers arising from the use of AI chatbots. However, the CFPB has been noticeably less clear on important questions related to the testing of AI/ML models for lending to mitigate discrimination, including disparate impact.

For example, the CFPB has signaled that it expects lenders to provide clear, accurate, and specific reasons for adverse credit decisions. It published a circular (2022-03) on adverse action requirements for explaining adverse credit decisions derived from algorithmic underwriting in May 2022[4] and another (2023-03) requiring creditors to provide accurate and specific reasons for denials in September 2023.[5]

The use of AI has been an element of the CFPB's work in other areas as well. The CFPB published an interpretive rule on the use of AI by digital marketers in August 2022,[6] and a spotlight on the use of AI chatbots in June 2023.[7] In an April 2023 guidance, the CFPB provided an analytical framework for how it would address abusive conduct and noted that its prohibitions could cover service providers who use an algorithm in ways that are abusive, including if the algorithm is used in marketing.[8] In June 2023, the CFPB proposed a rule for the use of automated valuation models to address the use of algorithms in assessments of the value of real estate (but not appraisals).[9] Lastly, the CFPB participated in an

---

[4] "Consumer Financial Protection Circular 2022-03: Adverse Action Notification Requirements in Connection with Credit Decisions Based on Complex Algorithms." CFPB, May 26, 2022. https://www.consumerfinance.gov/compliance/circulars/circular-2022-03-adverse-action-notification-requirements-in-connection-with-credit-decisions-based-on-complex-algorithms/.

[5] "Consumer Financial Protection Circular 2023-03: Adverse Action Notification Requirements and the Proper Use of the CFPB's Sample Forms Provided in Regulation B." CFPB, September 19, 2023. https://www.consumerfinance.gov/compliance/circulars/circular-2023-03-adverse-action-notification-requirements-and-the-proper-use-of-the-cfpbs-sample-forms-provided-in-regulation-b/.

[6] Interpretative Rule: Limited Applicability of Consumer Financial Protection Act's "Time or Space" Exception with Respect to Digital Marketing Providers. CFPB, August 10, 2022.https://files.consumerfinance.gov/f/documents/cfpb_time-or-space_interpretive-rule_signed_2022-08.pdf.

[7] Issue Spotlight: Chatbots in consumer finance. CFPB, June 6, 2023. https://www.consumerfinance.gov/data-research/research-reports/chatbots-in-consumer-finance/chatbots-in-consumer-finance/.

[8] "CFPB Issues Guidance to Address Abusive Conduct in Consumer Financial Markets." CFPB, April 3, 2023. https://www.consumerfinance.gov/about-us/newsroom/cfpb-issues-guidance-to-address-abusive-conduct-in-consumer-financial-markets/.

[9] "Proposed Rule: Quality Control Standards for Automated Valuation Models." CFPB, June 1, 2023. https://www.consumerfinance.gov/rules-policy/rules-under-development/quality-control-standards-for-automated-valuation-models/.

interagency statement outlining a commitment to enforce its authority to prevent discrimination and enforce existing consumer financial protection laws when creditors use automated systems.[10]

ii. Lack of guidance on identifying and mitigating discrimination in ML models

While the CFPB has been more explicit about appropriate uses of AI for marketing and real estate valuation, as well as the need for underwriting models to be explainable,[11] it has yet to clearly and concretely address critical issues such as what obligations lenders have to address the risk of discrimination and disparate impact, how they can identify and measure disparate impact, when they should search for alternative models, and how they can establish and maintain robust compliance management systems with these goals in mind. The CFPB has made brief, ad hoc oral statements on this topic in public appearances, noting the general expectation for lenders to search for LDAs when they use AI tools for underwriting. However, to date, the CFPB has not provided formal written guidance on this topic.

<u>We call on the CFPB to clarify the obligation of lenders to search for and implement less discriminatory alternatives to mitigate discrimination, and disparate impact in particular, and to provide ongoing guidance on how lenders should undertake this activity.</u>

iii. Negative impacts arising from lack of regulatory clarity

Absent clear regulatory guidance and supervisory expectations, participants using AI will be largely limited to select groups: (1) large financial institutions with resources to build models and conduct testing on their own, (2) bank and non-bank disruptors who will likely proceed forward without due regard for the potential of harm to consumers, (3) banks who conduct partnerships with fintechs, and (4) tech firms that contract with depositories to provide proprietary models. Left out will be many smaller community-focused banks and credit unions who often focus on reaching underserved consumers, as well as other depository institutions that feel hesitant to invest resources in technologies where compliance risk remains uncertain.

Among the groups using AI/ML models, a select few may build accurate models and use effective methods to search for LDAs. But it is likely that many will not search for LDAs, choosing to believe in "fairness through unawareness," while others will test but not in a robust manner. Some institutions may say their testing is more than fair when, in fact, they may actually be resistant to deploying best practices. In other cases, institutions may undertake genuine efforts but use methods that are ineffective or out of date, as there are legitimate questions regarding how best to test for and implement LDAs. As long as financial institutions fail to implement effective testing, their use of AI and ML will pose a serious risk to consumers. This could also have ramifications for competition, as it could mean that compliance costs are different and higher for good actors.

We have also heard anecdotally that there are internal conflicts within some financial institutions, with those internal champions who support searching for LDAs being prevented from doing so by others who fear that searching for LDAs will expose the institution to compliance risk (presumably given the

---

[10] "CFPB and Federal Partners Confirm Automated Systems and Advanced Technology Not an Excuse for Lawbreaking Behavior." CFPB, April 25, 2023. https://www.consumerfinance.gov/about-us/newsroom/cfpb-federal-partners-confirm-automated-systems-advanced-technology-not-an-excuse-for-lawbreaking-behavior/.

[11] "Director Chopra's Prepared Remarks on the Interagency Enforcement Policy Statement on 'Artificial Intelligence.'" CFPB, April 25, 2023. https://www.consumerfinance.gov/about-us/newsroom/director-chopra-prepared-remarks-on-interagency-enforcement-policy-statement-artificial-intelligence/.

uncertainty about how it should be done appropriately). This type of situation is the opposite of ideal. At the same time, these internal champions present an opportunity that could be capitalized upon with proactive guidance.

Providing clear guidance on supervisory expectations and best practices can move each of these groups toward pro-consumer, anti-discriminatory approaches. The difference between general statements calling for LDAs and concrete expectations and examples is important. Absent the latter, a door is left open for bad actors to claim their efforts – however insincere – are sufficient. But if concrete examples of how different institutions have implemented policies for LDA searches and testing exist – such as in supervisory highlights – these claims of innocence are greatly discredited, while well-intentioned institutions will also be enabled to conduct more robust testing.

iv. Objections to regulatory clarity are not well-founded

We acknowledge that alternative views exist that would reject our call for regulatory clarity. Some stakeholders have taken the perspective that clarity will tempt some actors to "game" regulations and that overly prescriptive instructions will put the CFPB in the position of picking "winners and losers" in financial markets. True, it is possible to be overly prescriptive, and the policy of a regulatory agency should never support the services of a single vendor over others. But these concerns can be addressed in any guidance issued by the CFPB, which can be designed to be flexible, not overly prescriptive, and agnostic to any particular method, while still providing necessary clarity and robust protections for consumers. Gaming seems problematic, but it is hard to understand how clarity on the need to proactively search for LDAs supports the evasion of regulation.

We also reject the idea that guidance and enforcement are mutually exclusive. To be certain, providing guidance is not a substitute for using enforcement authority to punish bad actors, but it is an important tool to provide the certainty needed for more well-intentioned actors to move forward during this time of rapid change and uncertainty. We believe clarity and corrective action are complementary; each fulfills a separate and equally necessary component of consumer financial protection frameworks. And, as mentioned above, the existence of examples that show the degree of commitment and proper procedures necessary to meet supervisory expectations becomes important to provide less basis for bad actors to respond to enforcement orders with litigation.

Algorithmic systems pose risks of discrimination

The risk of algorithmic systems resulting in biased outcomes that perpetuate and even exacerbate existing societal biases has been well-established in a wide range of research across multiple sectors.[12] Algorithmic discrimination occurs when an automated decision system repeatedly creates unfair or inaccurate outcomes for a protected class. While the risk of discrimination exists with traditional models, it is exacerbated by ML techniques for automated decision-making that rely on processing vast amounts of data using often opaque models.

Discriminatory results can arise from many sources, including underlying data and model design. Unrepresentative, incorrect, or incomplete training data, as well as biased data collection methods, can

[12] For example, see Barocas, Solon and Andrew D. Selbst. "Big Data's Disparate Impact." 104 California Law Review 671, 2016, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899; O'Neil, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Penguin Books, 2016; and Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. "Machine Bias." ProPublica, 2016, available at https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

lead to poor outcomes in algorithmic decision-making for certain groups. For example, researchers examining mortgage data found that one of the underlying reasons for differences in mortgage approvals between minority and majority groups is the limited credit history data on mortgages for minorities and low-income groups, leading to less precise predictions for such groups.[13] Data may also reflect historical biases, particularly the types of data sources used for underwriting, which are tainted by past discriminatory practices. Due to systemic racism, Black and Latinx Americans are more likely to have damaged credit or lower credit scores compared to their white counterparts[14] and are more likely to be sold high-cost, unmanageable loans.[15]

Biases can also be embedded into models through the design process, such as via improper use of protected characteristics directly or through proxies. With complex ML models utilizing hundreds or thousands of input features, chosen features may be proxies for protected characteristics. Choices made during the model development process can also affect its predictiveness regarding particular populations. The issue of potential discrimination is further compounded by the lack of transparency of complex ML models.

Biases in automated systems can result in incorrect or inaccurate decisions, leading to real harm to consumers due to financial exclusion or unfair pricing. The potential for automated systems to produce biased outcomes and "automate discrimination" has already been recognized by the CFPB and other civil rights agencies in an April 2023 joint statement, which also highlighted a commitment to enforce respective laws and regulations to address these very issues arising from the use of new technologies.[16] However, the CFPB has yet to take action directly addressing testing for and mitigating discrimination in automated systems, despite recent policy actions on other related AI issues.

The obligation to search for and implement LDAs builds off of existing disparate impact doctrine

Discrimination is not a new issue, and there is a long-standing doctrine aimed at tackling it. In particular, Title VII of the Civil Rights Act, the Fair Housing Act, and ECOA all provide a core legal foundation to combat discrimination and increase fairness in employment, housing, and credit, respectively. The legal regimes for all three laws have evolved to prohibit disparate treatment and disparate impact.

Disparate treatment involves intentionally treating different classes of people differently based on protected characteristics (e.g. race, gender, religion). In contrast, disparate impact involves facially neutral policies that have a disproportionately adverse effect on a protected class, regardless of intent. Such policies are prohibited if they either do not advance a legitimate business need or, if they do, an LDA exists that advances the same need. Disparate impact is particularly relevant to credit underwriting and

[13] Blattner, Laura & Scott Nelson. "How Costly is Noise? Data and Disparities in Consumer Credit." Papers 2105.07554, arXiv.org, 2021. https://arxiv.org/pdf/2105.07554.pdf.

[14] "Past Imperfect: How Credit Scores and Other Analytics 'Bake In' and Perpetuate Past Discrimination." National Consumer Law Center, May 2016. https://www.nclc.org/resources/past-imperfect-how-credit-scores-and-other-analyticsbake-in-and-perpetuate-past-discrimination/.

[15] Marte, Jonnelle. "Wells Fargo steered blacks and Latinos toward costlier mortgages, Philadelphia lawsuit alleges." Los Angeles Times, May 16, 2017. https://www.latimes.com/business/la-fi-wells-fargo-philadelphia-20170516-story.html.

[16] Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems. CFPB, Department of Justice, Equal Employment Opportunity Commission, and Federal Trade Commission, April 2023. https://files.consumerfinance.gov/f/documents/cfpb_joint-statement-enforcement-against-discrimination-bias-automated-systems_2023-04.pdf.

pricing, given historical biases embedded in data used for underwriting and pricing. It also has relevance in evaluating the fairness of marketing, servicing, and other financial practices.

Disparate impact doctrine has been clearly established in case law, regulatory guidance, and enforcement actions. For example, the CFPB explicitly states in Regulation B, the implementing regulation of ECOA, that the "effects test" concept is applicable to a creditor's determination of creditworthiness.[17] The official interpretations accompanying Regulation B further note that a creditor's practice may be prohibited if it is "discriminatory in effect because it has a disproportionately negative impact on a prohibited basis, even though the creditor has no intent to discriminate and the practice appears neutral on its face unless the creditor practice meets a legitimate business need that cannot reasonably be achieved as well by means that are less disparate in their impact."[18]

i. The concept of less discriminatory alternatives is already well-established in existing disparate impact doctrine.

Disparate impact doctrine typically involves a three-step test: (1) a plaintiff must establish a prima facie case showing that a policy or practice has a disparate impact on a disadvantaged group; (2) the defendant then has the burden of demonstrating a legitimate business justification for the practice; and (3) even where defendants provide a business justification, they can still face liability if there is an alternative approach that serves the same business needs with less disparate impact.

Multiple regulators have affirmed the obligation to search for LDAs as a core element of disparate impact doctrine. For example, the U.S. Department of Housing and Urban Development (HUD)'s discriminatory effects rule codified long-standing case law on discriminatory effects doctrine under the FHA, including noting that plaintiffs can still prevail in cases where defendants demonstrate a legitimate purpose for a challenged practice if they can show that this purpose "could be served by another practice that has a less discriminatory effect."[19] Similarly, the CFPB's official interpretation of Regulation B cited above incorporates the concept of less discriminatory alternatives.

Unfortunately, demonstrating the existence of a less discriminatory alternative can be challenging for plaintiffs in practice and is even more difficult in the context of complex, proprietary "black box" ML models. However, as discussed in the following section, it has become much easier for creditors to test for disparate impact and identify LDAs, particularly when employing ML.

ii. With the advent of ML models, finding LDAs has become much more feasible

Due to the unique capacity of ML models to improve through rapid iteration, it is now much less burdensome or time-intensive to search for and implement LDAs. For example, model multiplicity refers to a phenomenon identified in recent computer science and statistics research that shows that there are multiple possible models that are equally effective at a given task. "As a result, when an algorithmic system displays a disparate impact, model multiplicity suggests that other models exist that perform equally well but have less discriminatory effects. In other words, in almost all cases, *a less discriminatory*

[17] 12 Code of Federal Regulations § 1002.6 Rules concerning evaluation of applications

[18] Comment for 1002.6 - Rules Concerning Evaluation of Applications in Supplement I to Part 1002 - Official Interpretations. CFPB. https://www.consumerfinance.gov/rules-policy/regulations/1002/interp-6/.

[19] 24 Code of Federal Regulations § 100.500 Discriminatory effect prohibited

algorithm exists."[20] Model multiplicity implies that there no longer needs to be a significant tradeoff between performance and fairness.

Therefore, particularly in the context of ML models, it is no longer a question of whether or not a less discriminatory alternative that meets legitimate business needs can be found – this threshold question has already been answered in the affirmative. For example, the Equal Employment Opportunity Commission (EEOC) has noted that "one advantage of algorithmic decision-making tools is that the process of developing the tool may itself produce a variety of comparably effective alternative algorithms. Failure to adopt a less discriminatory alternative that was considered during the development process may give rise to liability."[21] Instead, the question now turns to how companies should go about finding and implementing LDAs, particularly in the case of ML models.

The CFPB should recognize the hurdle faced by plaintiffs seeking to meet the third prong for proving a disparate impact claim. A precondition for changing a "black box" is to first understand its logic. In this context, plaintiffs may find it impossible to demonstrate the existence of an LDA to an algorithmic model.

Model multiplicity facilitates model searches that can simultaneously improve performance and fairness. In these searches, a question arises about which goal is more important. We are concerned that the tradeoff, as there is, between business interests and fairness may swing too strongly toward the former. Rather, it is our belief that fairness can and should take on enhanced importance.

iii. Companies, not plaintiffs, are in a natural position to efficiently and effectively search for LDAs when employing ML models

<u>Requiring companies to proactively mitigate disparate impact by searching for and implementing LDAs is a natural and logical evolution of existing disparate impact doctrine and has significant potential to achieve anti-discriminatory policy objectives more effectively</u>. Particularly when it comes to ML models, the development process already involves weighing a series of choices and making continual refinements to optimize the performance of the resulting model, selected from a universe of potential models. Incorporating disparate impact as an additional lens to apply could be easily integrated into the typical model development process, leading to the selection of a model that will advance legitimate business interests but with less disparate impact.

A clear obligation to mitigate disparate impact and to search for and implement LDAs is also in line with broader emerging approaches towards rights-impacting AI (which includes decisions relating to insurance and credit, among others). For example, a recent memo from the Office of Management and Budget (OMB) on the governance of agency use of AI states that agencies must mitigate algorithmic discrimination when it is present before using AI, including mitigating disparities that lead to or perpetuate harmful bias or decrease equity.[22] This approach also aligns with the CFPB's general position

[20] Black, Emily, John Logan Koepke, Pauline T. Kim, Solon Barocas, and Mingwei Hsu. "Less Discriminatory Algorithms (October 2, 2023)." Georgetown Law Journal, Vol. 113, No. 1, 2024, Washington University in St. Louis Legal Studies Research Paper Forthcoming. https://ssrn.com/abstract=4590481.

[21] "Select Issues: Assessing Adverse Impact in Software, Algorithms, and Artificial Intelligence Used in Employment Selection Procedures Under Title VII of the Civil Rights Act of 1964." EEOC, May 18, 2023. https://www.eeoc.gov/laws/guidance/select-issues-assessing-adverse-impact-software-algorithms-and-artificial.

[22] Memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. OMB, March 28, 2024. https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf.

on the importance of establishing and pursuing strong compliance management expectations for all federal financial consumer protection laws.

The CFPB should clarify that there is a clear obligation under ECOA[23] to mitigate algorithmic discrimination and to search for and implement LDAs

The CFPB has already clearly stated that lenders are accountable to follow all anti-discrimination laws. The CFPB should provide regulatory clarity that the practice of searching for and implementing LDAs is part of the compliance expectations with existing anti-discrimination laws. CFPB staff have supported the expectation to search for LDAs as part of fair lending compliance during public events and indicated the CFPB's intention to require it. However, no official textual language clearly implements this standard. Given the need for timely action, the CFPB should go further and utilize a strong but expedient instrument to clarify this regulatory expectation.

The CFPB should lead the financial markets in a new direction where mitigating bias and searching for LDAs becomes the normal course of business. As noted previously, we are concerned by the inconsistency across the market regarding undertaking a genuine and robust search for LDAs. This concern will not be remedied without proper motivation from regulators. Relying on enforcement action to compel robust LDA searches will allow for harm to have already occurred and would also be limited to individual companies. Clarity from the CFPB will help to ensure consistent and robust efforts are undertaken across the entire market while also enabling the broader adoption of ML technologies by community-oriented institutions.

An advisory opinion (AO) could be leveraged to provide the necessary regulatory clarity on this issue. For example, in 2020, the CFPB published an advisory opinion on the use of special-purpose credit programs (SPCPs) by for-profit organizations. The 2020 AO provided clarity, expressed supervisors' expectations, and reinforced the CFPB's view that new rules were unnecessary. It is a model for fulfilling the purpose we propose in this letter for LDA searches.[24] The AO followed earlier messaging in Supervisory Highlights in November 2016 and a blog in July 2020. The edition in Supervisory Highlights called attention to successful SPCPs that provided credit to small businesses and mortgage applicants. The issue of Supervisory Highlights detailed how some lenders had provided mortgages at rates and terms below prevailing thresholds and in distressed neighborhoods.[25] The 2016 blog clarified commentary in Regulation B, permitting creditors to use affirmative advertising for SPCPs.[26] The AO's underlying message was that lenders could no longer hide behind claims of regulatory uncertainty to delay developing SPCPs.

---

[23] We note that an affirmative obligation to mitigate bias and to search for and implement LDAs may also arise from the prohibition of unfair, deceptive, and abusive acts and practices (UDAAP) under the Consumer Financial Protection Act (CFPA). Therefore, this obligation applies to models used for a broader range of financial products (beyond credit) and across various stages of the customer lifecycle (i.e., advertising, servicing, fraud, etc.).

[24] Consumer Financial Protection Bureau. "Consumer Financial Protection Bureau Issues Advisory Opinion to Help Expand Fair, Equitable, and Nondiscriminatory Access to Credit." Consumer Financial Protection Bureau, December 21, 2020. https://files.consumerfinance.gov/f/documents/cfpb_advisory-opinion_special-purpose-credit-program_2020-12.pdf.

[25] "Supervisory Highlights." Consumer Financial Protection Bureau, Summer 2016. https://files.consumerfinance.gov/f/documents/Supervisory_Highlights_Issue_12.pdf.

[26] Bernard, Susan M., and Patrice Alexander Ficklin. "Expanding Access to Credit to Underserved Communities." *Consumer Financial Protection* Bureau (blog), July 31, 2020. https://www.consumerfinance.gov/about-us/blog/expanding-access-credit-underserved-communities/.

The CFPB should provide financial institutions with guidance on how to search for and implement LDAs

To effectively operationalize the obligation to search for LDAs and compel lenders to take concrete action, the CFPB should provide guidance on how and when to undertake such a search. We acknowledge that a full set of tools and techniques to search for and implement LDAs are still in the development stage. Nevertheless, guidance from the CFPB will be crucial to work towards the development of consistent, robust approaches utilized across the market.

Methodological guidance should not be overly prescriptive, but to stimulate widespread adoption, the CFPB should still provide financial institutions with greater certainty and detail. Given the pace of evolution in this space, the CFPB should strategically leverage a mix of tools and approaches, such as interpretive rules, operating circulars, supervisory highlights, and research publications, to provide guidance and shape emerging best practices over time as they evolve.

Supervisory guidance, elaboration, and examples would be particularly useful in the following areas: what the obligation to mitigate bias and to search for and implement LDAs entails, operational aspects of developing and comparing LDAs, and appropriate metrics for fairness when testing for disparate impact.

i. The CFPB should establish clear expectations regarding what the obligation to mitigate disparate impact and to search for and implement LDAs entails

Generally, there are three stages identified in algorithmic modeling: (1) data input, (2) model development and evaluation, and (3) model deployment and ongoing monitoring. During the data input stage, modelers collect datasets, clean and process data, and identify significant variables. During model development and evaluation, a model is proposed, improved iteratively through machine learning, and evaluated. Finally, the model is deployed and monitored.

The obligation to mitigate disparate impact should entail companies taking reasonable, proactive steps during each stage of the model development pipeline. Employing less discriminatory approaches to mitigate disparate impact should be considered systematically and throughout the entire model development pipeline, including during the early stages, as limiting the search for LDAs only to later stages (or relying on post-processing techniques) can be less efficient and lead to sub-optimal outcomes.[27]

An emerging range of interventions can be utilized at each stage of the model development pipeline to mitigate disparate impact. For example, companies can evaluate the impact of different choices made during data pre-processing with respect to their impact on disparate impact. During the model development and evaluation stage, a range of techniques can be employed, including testing different combinations of features during feature selection, evaluating how different model types, model training approaches, and loss functions affect disparate impact metrics, hyperparameter tuning, etc.[28] The obligation to mitigate disparate impact through searching for and implementing LDAs should apply on an ongoing basis during the deployment of models, as well.

[27] For further details, see Black, Emily, Rakshit Naidu, Rayid Ghani, Kit T. Rodolfa, Daniel E. Ho, and Hoda Heidari. "Toward Operationalizing Pipeline-aware ML Fairness: A Research Agenda for Developing Practical Guidelines and Tools." ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, Oct 2023. https://arxiv.org/abs/2309.17337.

[28] For further examples, see Black, Emily, John Logan Koepke, Pauline T. Kim, Solon Barocas, and Mingwei Hsu. "Less Discriminatory Algorithms (October 2, 2023)." Georgetown Law Journal, Vol. 113, No. 1, 2024, Washington University in St. Louis Legal Studies Research Paper Forthcoming. https://ssrn.com/abstract=4590481.

The CFPB should provide broad guidance on appropriate steps that should be taken throughout the model development process to mitigate disparate impact. After first clarifying that LDA testing must be undertaken for regulatory compliance purposes, the CFPB should then elaborate on supervisory expectations. For example, lenders should be encouraged to take steps during the early stages of model development that mitigate disparate impact, conduct reviews for disparate impact during the model training phase as well as the deployment phase, and employ techniques to mitigate disparate impact by exploring LDAs.

Companies should be required to document their LDA search efforts for compliance purposes, including via model cards that include information on training data, performance metrics, fairness considerations, and mitigation strategies to address disparate impact risks.

ii. The CFPB should utilize a range of vehicles to highlight emerging best practices for how to search for and implement LDAs.

For companies to effectively search for LDAs, further clarification would be beneficial on operational aspects of LDA searches, particularly regarding: appropriate techniques for mitigating disparate impact; the frequency and level of effort of LDA searches expected of financial institutions; and considerations for a viable LDA.

Appropriate techniques for mitigating disparate impact: At the moment, there are varying views on what are considered to be best practices for developing LDAs. While rudimentary techniques exist to modify models to mitigate disparate impact (such as "drop-one" techniques),[29] a range of more advanced tools and techniques are emerging, including adversarial debiasing techniques, joint optimization, and Bayesian methods that use automated processes to more effectively search for modifications to reduce disparate impacts.[30] Additional techniques for identifying LDAs may include reweighting input features and adjusting decision thresholds.

Further development and robust testing of techniques will be needed to identify the most efficient and effective techniques that maximize reductions in disparate impact. For example, fewer techniques have been identified to date on interventions during the early stages of model development that are more likely to mitigate harm. Companies should be expected to utilize emerging good practices in terms of techniques for searching for LDAs.

Frequency of LDA search: Having model developers search for LDAs may be the most efficient means to address disparate impact, but to have the ideal effect when models and the outside world may change, it is not sufficient to test a model only upon its introduction. Establishing certain basic expectations regarding when and how frequently an LDA search should be undertaken will be necessary. For example, with respect to existing models, companies could be expected to undertake a search for LDAs on a regular

[29] NAACP Legal Defense and Educational Fund, Student Borrower Protection Center, and Upstart Network. "Fair Lending Monitorship of Upstart Network's Lending Model." Report of the Independent Monitor. Relman Colfax PLLC, November 10, 2021. https://www.relmanlaw.com/media/news/1182_PUBLIC%20Upstart%20Monitorship_2nd%20Report_FINAL.pdf.

[30] See Explainability & Fairness in Machine Lending for Credit Underwriting: Policy Analysis. FinRegLab, November 2023. https://finreglab.org/wp-content/uploads/2023/12/FinRegLab_2023-12-07_Research-Report_Explainability-and-Fairness-in-Machine-Learning-for-Credit-Underwriting_Policy-Analysis.pdf.

cadence as well as when significant changes to models or the economy[31] have occurred. For new models being developed, the search for LDAs should be integrated into the development process. The volume of applications may also factor into determining an appropriate cadence for LDA searches. For cases where fewer applications are made, it may take longer for a financial institution to collect enough new data to undertake a search for a new LDA.

Level of effort (LOE) of an LDA search: In addition to considering the frequency of LDA searches, a more intensive and higher LOE could reasonably be expected in certain situations. In particular, where levels of disparate impact remain high after re-testing, financial institutions should be expected to continue their LDA search. Moreover, creditors using more sophisticated ML techniques to build a model should also be expected to utilize more advanced techniques during the LDA search.

Choosing between "comparably" or "equally" effective as a standard for a viable LDA: The CFPB should also provide guidance on what financial institutions should consider when determining what is a viable LDA. In some cases, trade-offs may arise when fine-tuning models to mitigate disparate impact. In existing disparate impact doctrine, approaches vary on this topic. Some courts have stated that proposed alternatives must be "equally effective" or "equally valid" but not necessarily identical.

However, both the EEOC[32] and the Federal Housing Finance Authority (FHFA)[33] have either directly referenced or described in practice a less strict standard of "comparably effective." The Appendix to the Interagency Fair Lending Examination Procedures states that "if an alternative that is approximately equally effective is available that would cause a less severe adverse impact, the policy or criterion in question may constitute a violation."[34] "Comparably effective" would arguably be a more appropriate and consistent standard to be utilized for fair lending purposes, as it provides greater flexibility for situations where models may perform slightly less well but may significantly reduce disparate effects.

In practice, expectations regarding an appropriate standard may need to be fleshed out over time and will need to strike the right balance between fairness and any potential impacts on performance. While it may be premature for the CFPB to establish an explicit standard at this time, it should provide guidance and examples on how financial institutions should approach this issue. For example, where model developers have identified a standard range of model performance, any LDA that fits into this range should be deemed viable. The CFPB could also call attention to situations where financial institutions implemented "comparably effective" LDAs – a higher standard – to move actors forward where appropriate. Comparably effective LDAs may be more warranted where they result in significant reductions in disparate impact.

---

[31] Breeden, Joseph L, LiMing Brotcke, Bin Duan, Andy Johnson, and Charles Maner. "Impacts of COVID-19 on Model Risk Management." Model Risk Managers International Association, March 2021. https://mrmia.org/wp-content/uploads/2021/05/MRMIA-Impacts-of-COVID-19-on-Model-Risk-Management-v1.0-compressed.pdf.

[32] "Select Issues: Assessing Adverse Impact in Software, Algorithms, and Artificial Intelligence Used in Employment Selection Procedures Under Title VII of the Civil Rights Act of 1964." EEOC, May 18, 2023. https://www.eeoc.gov/laws/guidance/select-issues-assessing-adverse-impact-software-algorithms-and-artificial.

[33] AB 2021-04 on Enterprise Fair Lending and Fair Housing Compliance. Federal Housing Finance Agency, Dec. 20, 2021. https://www.fhfa.gov/SupervisionRegulation/AdvisoryBulletins/AdvisoryBulletinDocuments/AB%202021-04%20Enterprise%20Fair%20Lending%20and%20Fair%20Housing%20Compliance.pdf.

[34] Appendix to Interagency Fair Lending Examination Procedures. Federal Financial Institutions Examination Council, August 2009. https://www.ffiec.gov/pdf/fairappx.pdf.

On all of the above issues, good practices and techniques are still in the process of being developed. Therefore, the CFPB should leverage a wide range of formats to capture and elevate emerging good practices when searching for LDAs to shape industry behavior going forward and articulate supervisory expectations.

For example, as an expression of its supervisory authority, the CFPB regularly publishes "supervisory highlights" that identify problematic practices. These publications serve to proactively alter business practices going forward across the market. It is the exception to the rule when they call out a model approach, but short of that, highlights illustrating examples of financial institutions implementing effective model testing programs would be useful without being overly prescriptive.

The CFPB can also send signals to financial institutions on how to comply with Regulation B's expectations when using algorithmic modeling through publications from Research, Markets, and Regulations (RMR). For example, the CFPB could provide research similar to its 2014 paper on using Bayesian improved surname geocoding.[35]

The CFPB should leverage supervisory highlights and other similar tools on an ongoing basis to provide examples of successful testing regimes covering the above topics. A supervisory highlight achieves the goal of moving the market toward better outcomes. However, it does not implicate the CFPB in the politically untenable position of instructing lenders on how to underwrite. Conversely, it does not create a "safe harbor." Calling attention to successful examples of LDA searches provides an important signal to the market regarding supervisory expectations. A supervisory highlight could also provide a compliance officer with the support they need to overcome resistance to conducting LDAs.

iii. The CFPB should provide examples of appropriate methods that lenders should use to measure disparate impact.

Taking steps to mitigate bias and to search for and implement LDAs will require creditors to determine how to measure disparate impact. A range of measures could potentially be used to measure fairness. The choice of an appropriate fairness measure may differ across different products and different consumer segments, for example, when dealing with subprime versus prime products and customer segments.

In *Griggs v. Duke Power Co.*, where the Supreme Court first articulated the theory of disparate impact, it used an Adverse Impact Ratio (AIR)[36] to quantify the magnitude of the unfairness of Duke Power's hiring policies. Hiring is a binary decision. When the Court chose 80 percent as a threshold for unfairness, it established a baseline for using statistical measures to show how policies could be 'fair in form, but discriminatory in operation.'[37]

In addition to AIR, companies can test for unfairness in the price of loans by comparing interest rates for loans originated to protected and non-protected class members. A simple approach is to compare the

[35] Consumer Financial Protection Bureau. "Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity." Data and Research Reports, March 8, 2024. https://www.consumerfinance.gov/data-research/research-reports/using-publicly-available-information-to-proxy-for-unidentified-race-and-ethnicity/.

[36] An adverse impact ratio compares the outcomes experienced by a favored group with a non-favored "protected class" group. For an AIR on loan approval: (protected class approvals / applications) / (non-protected approvals / applications). For example, (3/6) / (5/6) = 60 percent.

[37] Griggs v. Duke Power Co., 401 U.S. 424 (1971), No. 124 (U.S. Supreme Court March 8, 1971).

means of interest rates for groups, but there may be reasons to use other variations of this general method.[38]

The use of AIRs and comparison of means methods raises important methodological questions. These questions concern the size of a data set and its demographic makeup. For example, what is considered a 'fair" AIR for a prime lender may differ greatly from that for a subprime lender. Because subprime lenders are likely to approve borrowers across a greater credit band and differences in mean credit scores exist between protected and non-protected class borrowers,[39] subprime lenders using credit scores for underwriting are likely to have higher (fairer) AIRs than prime lenders using similar inputs.

One point to acknowledge is that there are many ways to measure fairness. Comparing approval or interest rates of protected class members with white applicants is only one approach. It could be equally valid to review algorithms for performance, post-origination, to ensure that loans are made with an ability-to-repay standard.

This letter does not propose answers to the above questions but rather calls for the CFPB to provide guidance and examples on proper metrics and methodologies to measure fairness that capture emerging good practices and take context into consideration. Again, a range of vehicles, including supervisory highlights, can be used to share emerging best practices with the industry and clarify expectations regarding how to choose appropriate fairness metrics in different situations.

Conclusion

In conclusion, we urgently call on the CFPB to take concrete action to address algorithmic discrimination in credit underwriting and pricing by clearly articulating the expectation to search for LDAs under ECOA and providing guidance on what this search entails and how it should be conducted. Proactively searching for and implementing LDAs is a natural evolution of long-standing disparate impact doctrine, is particularly relevant in the ML context given model multiplicity, and has significant potential to more effectively mitigate disparate impact and address anti-discrimination objectives.

We acknowledge that some policymakers are skeptical of the need for explicit guidance and prefer to focus on supervisory and enforcement activities. We believe this approach is sub-optimal. By and large, financial institutions are currently not conducting robust LDA searches, either due to a lack of motivation or a lack of knowledge of appropriate tools and techniques. Regulatory clarity combined with sharing emerging best practices on how to conduct LDA searches can help to set a clear bar across the industry regarding fair and responsible practices and would actually strengthen and enable supervisory and enforcement activities. In addition, regulatory clarity will provide the comfort and certainty for more financial institutions to leverage complex ML models, which can benefit financial inclusion.

The CFPB should clarify the expectation under Regulation B that lenders conduct robust searches for LDAs when developing and deploying ML models to mitigate disparate impact. The CFPB should also utilize a range of tools to provide ongoing guidance and examples of emerging best practices to shape expectations on how the search for LDAs should be properly undertaken.

---

[38] Zhang, David Hao, and Paul Willen. "Do Lenders Still Discriminate? A Robust Approach for Assessing Differences in Menus." 2020, Federal Reserve Bank of Boston Research Department Working Papers No. 20-19. https://doi.org/10.29412/res.wp.2020.19.

[39] Martincheck, Kassandra, Alex Carther, and Breno Braga. "Credit Health during the COVID-19 Pandemic." Urban Institute, March 8, 2022. http://urbn.is/credit.

Of course, a broader range of actions will still be needed to fully address fair, responsible use of AI in the financial sector, such as with respect to biased data and appropriate data governance, internal and external transparency, and model risk management (MRM). For example, existing MRM guidance from the Federal Reserve SR 11-7 and OCC 2011-12 should be updated to account for the unique risks posed by AI, such as greater complexity, opacity, and potential for bias. Nonetheless, we believe the actions described in this letter can have prompt and concrete effects to help address algorithmic discrimination.

Sincerely,

Jennifer Chien
Senior Policy Counsel, Financial Fairness
Consumer Reports
jennifer.chien@consumer.org

Adam Rust
Director of Financial Services
Consumer Federation of America
arust@consumerfed.org